

ОБРАБОТКА ПОКАЗАНИЙ ТЕПЛОСЧЁТЧИКОВ СТАТИСТИЧЕСКИМИ ИНСТРУМЕНТАМИ В ПРОГРАММЕ STATISTICA

П.А. Зорин

Научный руководитель: О.В. Стукач, д.т.н., профессор ИШИТР
Томский политехнический университет
raz8@tpu.ru

Введение

Объем накопленных данных с теплосчётчиков, установленных в городе Томске, можно отнести к категории Больших данных. В настоящее время проведение анализа современными статистическими инструментами таких данных не является обязательным в сложившихся отраслевых стандартах и программном обеспечении. Потребителем такой информации могут быть сами конечные пользователи – жители и собственники многоквартирных жилых домов, которые на основании полученных выводов могут принять управленческие решения для организации первоочередных мероприятий по повышению энергоэффективности домов.

Описание алгоритма

Для анализа использованы среднесуточные показания теплосчётчиков за период с 1 декабря 2013 года до 30 сентября 2017 года. Архивы были дополнены данными среднесуточных температур наружного воздуха за измеряемый период. В первую очередь была произведена очистка данных от некорректных показаний, обусловленных либо нарушением технологического режима, либо неисправностью прибора учёта.

Схемы и типы нагрузки, попадающие под учёт теплосчётчика, варьируются. Например, под учёт может попадать отопление жилого дома с горячим водоснабжением, как открытого, так и закрытого типа. Некоторые теплосчётчики, при этом, учитывают только отопление. Распределение количества узлов учета по типам нагрузки показано в таблице 1. Чтобы привести данные к единому виду, необходимо вычислить тепло, затраченное на нужды ГВС и вычесть его из показаний прибора по формуле 1. Необходимо отметить, что вычисление $Q_{\text{ГВС}}$ производится расчётным методом, так как фактический учёт не ведётся.

$$Q_{\text{отопление}} = Q_{\text{приборное}} - Q_{\text{ГВС}} \quad (1)$$

Таблица 1. Распределение узлов учёта по типам нагрузки

Тип нагрузки под учёт	Количество узлов учёта
Отопление	108
Отопление+ГВС (открытая схема)	481
Отопление+ГВС (закрытая схема)	484

Далее, дома необходимо разбить на группы по

признакам. Это можно сделать вручную по фактическим признакам: этажность, материал стен, площадь. Но мы будем использовать инструменты кластерного анализа в пакете Statistica. Уже внутри полученных групп произведём дальнейшую обработку данных сравнением затрат на отопление у каждого дома внутри группы. Для этого для каждого дома найдены линейные коэффициенты регрессии, моделирующие поведение дома. Сравнивая коэффициенты уравнения можно ранжировать дома по шкале энергоэффективности.

Тестирование алгоритма группировки домов

При разбиении на группы гетерогенного множества домов с помощью кластерного анализа нужно учесть, что характеристики «схема учёта» и «материал стен» являются номинальными и не могут быть преобразованы в количественные или порядковые. Поэтому будут использованы количественные характеристики «Площадь дома», «Удельная тепловая характеристика» и порядковая характеристика «Этажность».

Для нахождения кластеров был использован метод Уорда. В этом методе в качестве целевой функции применяют внутригрупповую сумму квадратов отклонений, которая есть не что иное, как сумма квадратов расстояний между каждой точкой (объектом) и средней по кластеру, содержащему этот объект. На каждом шаге объединяются такие два кластера, которые приводят к минимальному увеличению целевой функции, т.е. внутригрупповой суммы квадратов отклонений.

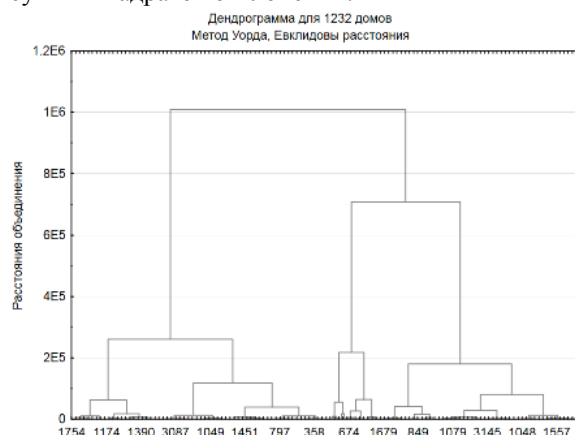


Рис. 1. Дендрограмма разбиения на кластеры

После оценки дендрограммы, приведённой на рисунке 1 было выбрано оптимальное количество кластеров – 4. Кластеры были найдены методом k -средних, затем были построены графики координат центров для каждого кластера, представленного на

рисунке 2.

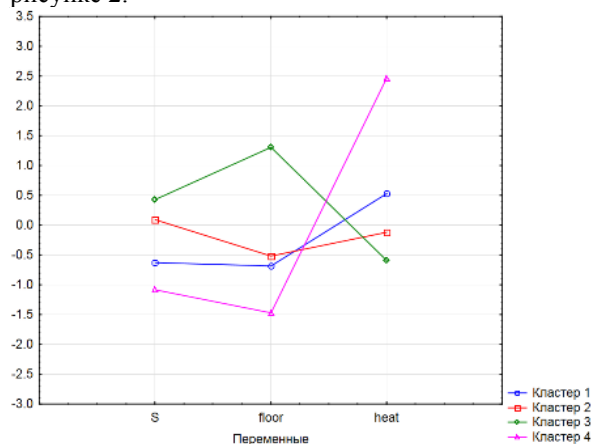


Рис. 2. Графики координат центров кластеров

Анализируя график координат центров необходимо отметить, что центры разных кластеров существенно различаются для каждой переменной (S – площадь, floor – этажность, heat – удельная тепловая характеристика), зафиксировав индивидуальные свойства каждого кластера.

Критерии качественной оценки энергоэффективности домов

Для получения математических моделей домов, имеющих сходные характеристики и входящих в один кластер, были вычислены коэффициенты теплоотдачи для каждого случая измерения:

$$G = \frac{0,001 \cdot P}{(T_{\text{вн}} - T_{\text{нар}}) \cdot S} \cdot \frac{\text{Вт}}{\text{м}^2 \cdot \text{Гр}} \quad (2)$$

Где: P – потребленное тепло за сутки по данным теплосчётчика, пересчитанное в кВт;

$T_{\text{вн}}$ – температура комфорта внутри жилых помещений, взята в виде константы 24° C;

$T_{\text{нар}}$ – среднесуточная температура наружного воздуха в данные сутки.

Далее, использованы данные теплосчётчиков двух похожих домов, вошедших в кластер №1, характеристики приведены в таблице 2.

Таблица 2. Описание характеристик домов

№	Адрес	Материал	S, м ²	Этажность	Число измерений
1	Вершинина 24/2	Кирпич	1226.4	4	827
2	Кузнецова 28	кирпич	1242.6	3	691

Для примера построены диаграммы рассеяния на рисунках 3 и 4 зависимости G от теплового напора с получением коэффициентов регрессии вида:

$$f(x) = a \cdot x + b \quad (3)$$

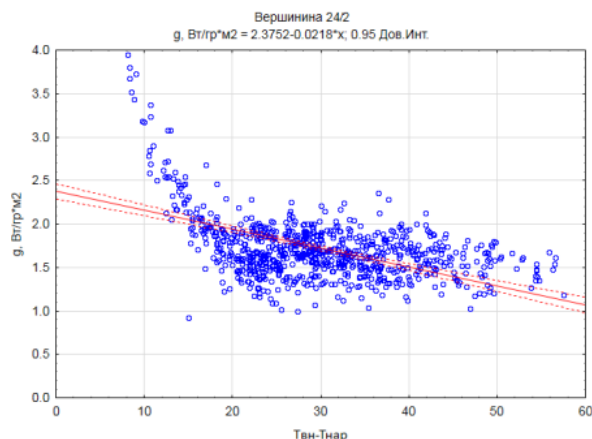


Рис. 3. Диаграмма рассеяния для дома №1

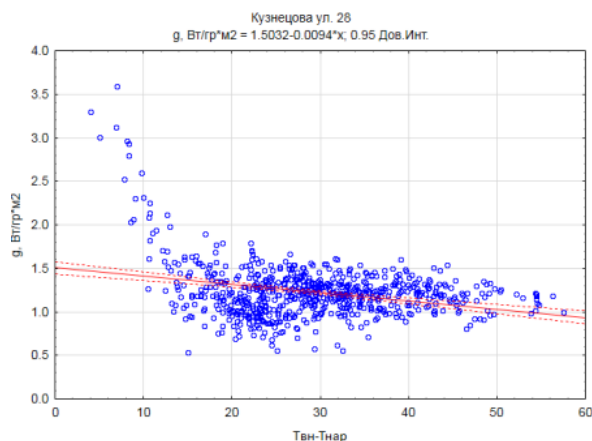


Рис. 4. Диаграмма рассеяния для дома №2

Заключение

Анализируя полученные математические модели энергоэффективности двух домов, можно понять, что, несмотря на схожесть их свойств – площадь и материал стен, ведут они себя по-разному, дом Кузнецова ул. 28 более эффективен, чем дом Вершинина 24/2. На основе моделей по группам можно найти дома наиболее энергоэффективные. Простейшее обследование позволит сделать выводы для принятия решений о повышении энергоэффективности для остальных домов.

Список использованных источников

1. Ершов И.А., Стукач О.В. Использование корректного нормирования статистических данных в кластерном анализе // Современные технологии поддержки принятия решений в экономике: сб. тр. Всеросс. научно-практ. конф. студентов, аспирантов и молодых ученых. г. Юрга, 19-20 ноября 2015 г. - ТПУ, 2015. - С. 32-34.
2. Стукач О.В. Программный комплекс STATIS-TICA в решении задач управления качеством – ТПУ, 2014 – 163 с.
3. Халафян А.А. Статистический анализ данных. – М.: Бином, 2005. – 512 с.